DOCUMENT RESUME

ED 340 782 TM 018 038

AUTHOR Archbald, Douglas A.; Porter, Andrew C.
TITLE A Retrospective and an Analysis of Roles of Mandated Testing in Education Reform.
SPONS AGENCY Congress of the U.S., Washington, D.C. Office of Technology Assessment.
PUB DATE 16 Jan 90
NOTE 79p.; Contractor report prepared for the Office of Technology Assessment titled "Testing in American Schools: Asking the Right Questions." For related document, see TM 018 025.
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Accountability; Curriculum Development; *Educational Change; Educational History; Educational Objectives; Elementary Secondary Education; Evaluation Methods; *Evaluation Utilization; Multiple Choice Tests; *Policy Formation; School Districts; *State Programs; Student Evaluation; Testing Problems; Testing Programs; *Test Use
IDENTIFIERS Large Scale Programs; *Mandated Tests

ABSTRACT

The role and influence of mandated testing in educational reform are reviewed. Mandated testing refers to large-scale (districtwide or statewide) multiple-choice testing programs used for policy purposes of evaluation and accountability, which includes nationally normed standardized achievement tests and tests custom-developed to reflect state and district educational objectives. Part 1 of this document provides a historical context dating back to the mid-1970s, describing the growth of mandated testing and its roles. Part 2 discusses current critiques of mandated testing and disillusionment with the top-down model. The detrimental effects of mandated testing on curriculum and teaching are reviewed, with the suggestion that criticism goes beyond evidence for too much testing and central control. Part 3 describes the growing support for comprehensive curriculum reform and for more coherent approaches to education reform in general. It is argued that testing and restructuring should be driven by a single, curriculum-based set of goals. Four tables provide detailed information, and two appendices give supplemental information. (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED340782

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

# A Retrospective and an Analysis of Roles of Mandated Testing in Education Reform

by

Douglas A. Archbald
Assistant Professor
Department of Educational Development
University of Delaware

Andrew C. Porter
Director
Wisconsin Center for Education Research
University of Wisconsin-Madison

January 16, 1990

Prepared for the Congressional Office of Technology Assessment, Washington, D.C. We wish to thank David Kirp, Edward Roeber, and Bruce Wilson for their helpful comments on an earlier draft of this report.

This contractor document was prepared for the OTA assessment entitled *Testing in American Schools: Asking the Right Questions*. This is being made available because it contains much useful information beyond that used in the OTA report. However, it is not endorsed by OTA, nor has it been reviewed by the Technology Assessment Board. References to this should cite the contractor, not OTA, as the author.

TM 018038



BEST COPY AVAILABLE

## INTRODUCTION

"There's an old Army saying, 'If it moves, salute it,'" said Gregory Anrig, president of the Educational Testing Service. "Today, some reformers seem to be saying 'If it moves, test it.'" So began Edward Fiske's 1988 article in the New York Times, "America's Test Mania."

Testing has been a dominant tool of education reform policy since the 70s. Tests have potent symbolic appeal.[1] Simple and understandable, they signify quality control and accountability to legislators and the public.[2] Testing is advocated to restore high educational standards and accountability for performance, expose poorly functioning schools and incompetent practitioners, prevent social promotion and the granting of unearned diplomas, and diagnose curriculum deficiencies and individual students' learning needs. Testing is a billion dollar industry and a central component of educational policy in virtually every state.

This report is about mandated educational testing of students and its

---

[1]Gene Glass and Mary C. Ellwein, "Reform By Raising Test Standards," CRESST Evaluation Comment (UCLA Center for Studies in Evaluation), December, 1986, pp. 1-6.

[2]Peter Airasian, "State Mandated Testing and Educational Reform: Context and Consequences." American Journal of Education, May, 1987, pp. 394-412. According to a 1987 Gallup survey, 70% of adults support public comparisons of schools on test scores, while 14% are opposed; 72% percent believe such comparisons can "encourage [low-performing] schools to try to do a better job." The 1986 Gallup survey revealed that 85% of respondents supported a state test requirement for teachers in the field in which they intend to teach.

role and influence in education reform. Although educational testing covers a broad range of assessment devices, our use of the term "mandated testing" refers to large scale (district- or state-wide) multiple-choice testing programs used for policy purposes of evaluation and accountability. This includes two broad categories of tests closely associated with education reform -- nationally normed standardized achievement tests and tests custom developed to reflect state and district educational objectives.

Part 1 provides a historical context dating back to the mid-70s, when minimum competency testing began on a large scale. We describe the growth of mandated testing and how goals of testing and state- and nationally-initiated, top-down education reform have fit together and conclude with a discussion of outcomes of these reforms and the role of mandated testing. Part 2 takes up the current critique of mandated testing and disillusionment with the top-down model. We focus on concerns over detrimental effects of mandated testing on curriculum and teaching, and suggest the critique goes well beyond the evidence. Beliefs that top-down reforms have produced too much testing and central control and too little success have lead to support for "bottom-up" approaches going under the heading of "restructuring." Part 3 describes the growing support for comprehensive curriculum reform and for more coherent approaches to education reform more generally. This view is critical of the lack a curriculum-based vision in education reform and argues reformers must design systems of policies unified by a clear and coherent conception of curriculum. Testing and restructuring should be driven by a single curriculum-based set of goals.

# PART 1

## MANDATED TESTING AND TOP-DOWN EDUCATION REFORM:

## HISTORY, RATIONALE, AND OUTCOMES

### INTRODUCTION

Education reform over the last 15 years has largely been top-down reform. The impetus has come from outside the education system: reform reports, business groups, state and national education leaders, and the media. Reformers proclaimed educational standards were lax,[3] there was too much misuse of time, too many teachers were incompetent,[4] and basic skills and core academic goals of schooling were receiving short shrift. Schools were blamed for social ills, joblessness, and declining economic productivity.[5] However, most crucial in declining confidence in public schools -- the pivotal factor -- was poor performance on standardized tests.[6] "Declining SAT scores" became a rallying slogan in the 70s and created a receptive audience for the early 80s

---

[3]See, for instance, Frank Armbruster Our Children's Crippled Future. (New York: Quadrangle Books, 1977).; Paul Copperman, The Literacy Hoax: The Decline of Reading, Writing and Learning in the Public Schools and What Can Be Done About It (NY: Morrow, 1978); Morris Cline, Why Johnny Can't Add (NY: Martin's Press, 1973).

[4]"Help? Teacher Can't Teach," Time, Vol. 16, June, 1980, pp. 54-63.

[5]The National Commission on Excellence in Education, A Nation At Risk: The Imperative for Educational Reform (Washington, DC: US Government Printing Office, 1983).

[6]George Madaus, "Testing and Policy -- True Love, Shot Gun Wedding or Marriage of Convenience?" Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, April, 1984.

reform reports.[7]

Mandated testing of students and top-down education reform have gone hand-in-hand over the last 15 years. During these years, many educational fads and issues have come and gone, but improved measurement of student and school performance for evaluation and accountability has been a consistent education reform priority. Cheap, easy, and accepted by the public, each successive wave of alarm and reform has brought increased testing.

The following descriptions of different reforms provide a brief history of mandated testing in the context of top-down education reform. These description show testing has continually expanded to measure a progressively broader range of schools' "outputs" and gradually acquired more influence through efforts to link stronger consequences to test performance. The pervading rationale has been one of evaluation and accountability: measuring the performance of students or schools to assess whether their performance meets expected standards. Whether standards have been called basic skills, literacy, or excellence, they have largely been defined by external authorities and measured by mandated tests.

---

[7]For an analysis of the quality of the documentation and data in the four major education reports issued in 1983 see, Lawrence Stedman and Marshall Smith, "Recent Reform Proposals for American Education." (Madison, WI: Wisconsin Center for Education Research, 1983.)

## THE MINIMUM COMPETENCY TESTING MOVEMENT

The 70s were a kind of transition period in testing. Prior to the 70s, the main form of externally developed test was the "off-the-shelf" nationally normed standardized test, such as the Iowa Test of Basic Skills. Schools did (and still do) aptitude testing and achievement testing at their discretion to diagnose childrens' learning needs, to place children in programs, and to evaluate programs. Federally-funded compensatory education in the 60s greatly expanded local use of standardized tests for student identification, placement, and evaluation purposes.

During the 70s, educational testing grew well beyond these local uses of standardized tests. While the 1983 reform reports and the wave of state education policies that followed represented a peak in education reform, rumblings of discontent with schools in the 70s produced the initial major expansion of state-mandated testing. This helped lay the foundation for the 80s expansion of testing following "A Nation At Risk." This report, "merely blew on fires already smoldering and combined with the results of several large research studies of schooling initiated in the 1970s, pushed education reform and renewal to center stage..."[8]

Minimum competency testing emerged in the conservative "back to the basics" movement following the 60s' free schools and student-centered learning. Airasian and Madaus write:

---

[8]Theodore Sizer, in the Preface (p. ix) of Thomas Timar and David Kirp, Managing Educational Excellence (NY: Falmer, 1988).

> Minimum competency testing for pupil certification was born of a number of perceptions: (1) that the pupils' basic skills of literacy and numeracy had seriously deteriorated; (2) that pupils were being promoted from grade to grade automatically, regardless of achievement; (3) that pupils did not have the necessary skills to survive in society; and (4) that the high school diploma had become devalued and meaningless as a credential (p. 107).[9]

Minimum competency tests have been used mainly to add an external standard to guide teachers' grade to grade student promotion decisions and to determine whether or not a student is entitled to a high school diploma. Students typically must achieve a preset passing score for grade promotion or to receive a diploma. Almost always, a number of re-takes are permitted.

Minimum competency tests are criterion-referenced, in that they measure performance in relation to specific pre-defined skills, such as basic vocabulary and reading comprehension, computational skills, and functional skills like understanding and filling out job application forms, interpreting bus schedules, and simple money management. (Norm-referenced tests reveal a student's relative performance in relation to the performance of other students, e.g., a percentile ranking; usually national norms are used.) While competency tests can be based on many different kinds of tasks (e.g., essays,

---

[9]Peter Airasian and George Madaus, "Linking Testing and Instruction: Policy Issues," Journal of Educational Measurement, vol. 20, No. 2, Summer, 1983, pp. 103 - 118.

speech, problem-solving), in practice the multiple-choice format is almost always used.

In addition to the functional nature of their content, two other features distinguish minimum competency tests from other forms of student testing. First, because minimum competency tests use specific passing scores, they require some form of standard setting procedure to determine and justify the "cutoff score."[10] Second, minimum competency tests are always administered on a census basis; that is, each student takes the test. Testing discussed later for assessing performance of programs, schools, districts, or groups of students can be administered on a sample basis.

Minimum competency tests are also viewed as means to insure quality control and create school-level accountability for the achievement of minimum functional skills. Many states and districts aggregate individual student scores to reveal school-level passing rates or average scores.

Prior to 1975 only a few states mandated competency testing; by 1983 over 30 did. As of 1987, 11 states required passing a minimum competency test

---

[10]Standard-setting is an entire field unto itself, with its own literature and complex legal, political, and technical issues. For discussion of this, see Robert Linn, George Madaus, and Joseph Pedulla, "Minimum Competency Testing: Cautions on the State of the Art," American Journal of Education, November, 1982, pp. 1-35; and, Richard Jaeger, "An Iterative Structured Judgment Process for Establishing Standards on Competency Tests: Theory and Application," Educational Evaluation and Policy Analysis, vol. 4, No. 4, Winter, 1982, pp. 461-475.

for graduation.[11] Between 1983 and 1987, 22 states initiated or expanded minimum competency testing.[12]

## THE GROWTH OF STATE ASSESSMENT PROGRAMS AND ACCOUNTABILITY TESTING

Minimum competency testing is limited to basic functional skills. Minimum competency tests are usually state tests or district tests, and usually do no allow comparisons to national norms. While mandated testing has been on the rise for several decades, the mid-80s wave of reform accelerated the expansion of mandated testing. How much? Amazingly, no one seems to know exactly.

There are few precise figures on trends in the frequency and in the time and costs expended for mandated testing -- but it clearly has increased in two basic ways. There have been increases in the scope and in the policy consequences of mandated testing as states have sought to hold schools accountable for a broader range of educational outcomes.

Between 1983 and 1987, the number of states requiring districts to test all students at multiple points in the K - 12 sequence increased from 38 to 44.[13] In the 1990 Carnegie Foundation national teacher survey (n = 21,389),

---

[11]Office of Technology Assessment, State Educational Testing Practices (Washington, DC: December, 1987, p. 99-100).

[12]Margaret Goertz, State Educational Standards in the 50 States: An Update (Princeton, NJ: Educational Testing Service, 1988).

[13]Margaret Goertz, op cit.

42% of teachers reported there has been "more use of externally developed tests to determine student promotion and graduation" at their school since 1983. (33% reported "same;" 4%, "less;" and 21%, "don't know.") (This is probably a conservative estimate of testing increases since the question asks only about promotion/graduation tests.)[14]

Tables 1 and 2 provide some figures concerning the scope of state mandated testing. Table 1, based on figures through 1986,[15] groups states into three categories, depending on the number of different grade levels at which basic skills tests (reading, mathematics, language arts) are administered. Eighteen states (including Washington DC) administer basic skills tests at 4 or more grade levels in the K - 12 sequence.

[ TABLE 1 HERE ]

---

[14]The Carnegie Foundation for the Advancement of Teaching, The Condition of Teaching (Princeton, NJ, 1990).

[15]Office of Technology Assessment, State Educational Testing Practices (Washington, DC: December 1987).

The data in Tables 1-3 are based on telephone interviews conducted by the Office of Technology Assessment. Because of ongoing changes in state testing policy, figures in the Tables 1 - 3 are already dated for a number of states. Also, state testing programs are difficult to classify and categorize in precise ways because these programs can vary on so many different dimensions. For instance, a test of "mathematics" can emphasize very different domains: basic computation, use of formulas of particular mathematics fields (e.g., trigonometry, geometry, algebra), applied mathematical reasoning, or various combination thereof.

State tests vary in other ways. They can be: referenced to different populations (e.g., state or national), "off-the-shelf" (e.g., Stanford Achievement Test) or "custom-developed" to match state curricula, census- or sample-based (and there are different sampling approaches), administered with varying degrees of flexibility at the local level, and require different degrees of aggregation and specificity in reporting at the local level.

TABLE 1 States Grouped by Total Number of Grades of Basic Skills Testing

| None | 1 - 3 | 4 + |
|---|---|---|
| IOWA | ALASKA | ALABAM |
| MASSACH | ARKANSA | ARIZONA |
| MONTANA | COLORAD | CALIFOR |
| NEBRASK | CONNECT | DC |
| NEVADA | FLORIDA | DELAWAR |
| NEWHAMP | IDAHO | GEORGIA |
| NEWJERS | ILLINOI | HAWAII |
| NORTHDA | LOUISIA | INDIANA |
| OHIO | MAINE | KANSAS |
| OKLAHOM | MARYLAN | KENTUCK |
| VERMONT | MICHIGA | MINNESO |
| WISCONS | MISSOUR | MISSISS |
| | NEWMEXI | NEWYORK |
| | OREGON | NORTHCA |
| | PENNSYL | RHODEIS |
| | SOUTHCA | TENNESS |
| | SOUTHDA | TEXAS |
| | UTAH | WESTVIR |
| | VIRGINI | |
| | WASHING | |
| | WYOMING | |

Standardized nationally-normed general achievement tests in reading and mathematics, at one time the modal form of testing, are increasingly being supplemented by testing programs in other subjects, like science or history. Between 1984 and 1988, the number of states with statewide achievement tests in mathematics grew from 34 to 43, and in science, from 13 to 28.[16]

Table 2 groups states by the number of different subjects they test. The modal category is 4 to 6 subjects, which includes (in virtually all cases) mathematics and reading, and two to four other subjects such as writing, science, history, civics, health, life skills, etc. Over the last 20 years, states have moved well beyond their traditional role of assessing basic skills. (States absent from Table 2 reported in 1986 having "no state-wide mandated testing program.")

[ TABLE 2 HERE ]

Also, the stakes of testing have increased. Business leaders and policy-makers often linked support for expensive reform packages to the willingness of state education agencies and school districts to accept public disclosure of test performance and other consequences that would strengthen accountability to the public. Thus, more states now produce public reports of

---

[16]Council of Chief State School Officers, CCSSO Assessment and Evaluation: Notebook, (Washington, DC: CCSSO, State Education Assessment Center, 1984); Council of Chief State School Officers, State Education Policies Related to Science and Mathematics, (Washington, DC: CCSSO, State Education Assessment Center, 1984).

TABLE 2 States Grouped by Total Number of Subjects Tested

| 1 - 3 | 4 - 6 | 7 - |
|---|---|---|
| ALASKA | ALABAM | CONNECT |
| ARKANSA | ARIZONA | DC |
| COLORAD | CALIFOR | HAWAII |
| FLORIDA | DELAWAR | MINNESO |
| KANSAS | GEORGIA | MISSISS |
| MARYLAN | IDAHO | NEWYORK |
| NEWMEXI | ILLINOI | |
| OREGON | INDIANA | |
| RHODEIS | KENTUCK | |
| TEXAS | LOUISIA | |
| WASHING | MAINE | |
| WYOMING | MICHIGA | |
| | MISSOUR | |
| | NORTHCA | |
| | PENNSYL | |
| | SOUTHCA | |
| | SOUTHDA | |
| | TENNESS | |
| | UTAH | |
| | VIRGINI | |
| | WESTVIR | |

school and district test results. More districts are doing the same with school results on their own testing programs.

Table 3 groups states according to whether they produce public reports that compare either districts or schools on results of state tests. Table 3 also includes information on the number of subjects the state testing program covers and the number of grades at which testing is conducted. (States absent from Table 3 reported in 1986 no mandated statewide testing program.)

[ TABLE 3 HERE ]

Another study reports 26 states link school-level performance data (primarily test results) to rewards and sanctions aimed at improving schools.[17] Among the strategies states have employed to increase the leverage of testing are cash bonuses for high performance, state awards, and regulatory waivers for high performance; and, for low performance, heightened regulation, forced effective schools planning, probationary accreditation status, and finally deregistration and state takeover -- what has been called "academic bankruptcy" legislation.

It should be pointed out that reforms linking achievement tests and accountability are not just a state phenomenon. Increases in state testing have been dramatic and well-documented -- it is much easier to collect and

---

[17]Council of Chief State School Officers, State Accountability Systems (Washington, DC, 1987). See also "Data Bank" column, Education Week, Oct 12, 1988, p. 14.

TABLE 3

| States Not Comparing Districts or Schools | | | States Comparing Districts or Schools | | |
|---|---|---|---|---|---|
| | No. Subj. | No. Grades | | No. Subj. | No. Grades |
| ARKANSA | 3 | 3 | ALABAM | 5 | 6 |
| COLORAD | 3 | 3 | ALASKA | 2 | 2 |
| DC | 7 | 5 | ARIZONA | 4 | 12 |
| IDAHO | 6 | 1 | CALIFOR | 5 | 4 |
| ILLINOI | 6 | 3 | CONNECT | 7 | 3 |
| KANSAS | 2 | 5 | DELAWAR | 6 | 9 |
| KENTUCK | 5 | 7 | FLORIDA | 3 | 3 |
| LOUISIA | 6 | 3 | GEORGIA | 5 | 5 |
| MICHIGA | 5 | 3 | HAWAII | 8 | 5 |
| MISSISS | 9 | 8 | INDIANA | 4 | 5 |
| MISSOUR | 4 | 2 | MAINE | 6 | 3 |
| NEWMEXI | 3 | 3 | MARYLAN | 3 | 3 |
| OREGON | 3 | 3 | MINNESO | 7 | 4 |
| RHODEIS | 3 | 4 | NEWYORK | 7 | 5 |
| SOUTHCA | 6 | 3 | NORTHCA | 5 | 5 |
| SOUTHDA | 5 | 3 | PENNSYL | 6 | 3 |
| UTAH | 4 | 2 | TENNESS | 6 | 8 |
| VIRGINI | 5 | 3 | TEXAS | 3 | 6 |
| WESTVIR | 5 | 4 | WASHING | 3 | 3 |
| WYOMING | 2 | 3 | | | |

report data from 50 state agencies than from thousands of school districts. However, testing programs administered by school district evaluation specialists and serving district evaluation and accountability purposes have also mushroomed in the last several decades. (With the growth of state testing, districts have in some cases curtailed their own testing programs or "piggy-backed" their testing on to state testing in a variety of ways.) No one knows precisely how many tests are taken by students each year, but virtually all medium and large size school systems in our country have testing programs with achievement testing, ability testing, and various forms of specialized diagnostic testing. Big city systems typically do the most testing.[18]

The National Center for Fair and Open Testing estimates K - 12 students in the U.S. take around 100 million standardized tests, about an average of 2.5 standardized tests per student per year (this includes state and district testing). The National Commission on Testing and Public Policy estimates standardized testing costs between $700 million to $900 million yearly in purchasing costs and administration time, or about $17 to $22 per student per year.

---

[18]For information on testing programs in urban districts, see Michael Casserly, "Statistical Profiles of the Great City Schools: 1970-82." (Washington, DC: Council of Great City Schools, 1983).

## RECENT ASSESSMENT REFORMS: PERFORMANCE ASSESSMENT AND INDICATOR SYSTEMS

### Performance Assessment

Performance assessment has attracted much attention recently because it promises assessment of a broader spectrum of outcomes and can produce more useful measures of performance for teachers. The more challenging forms of performance assessment require students to exercise creative thinking, to produce and integrate ideas, and to work at a problem for a sustained period of time (unlike the few seconds or minutes involved in answering a multiple choice question). Scoring of students' performance tasks is done by groups of teachers or other trained specialists. A student's performance (oral presentation, setting up a laboratory experiment) or product (e.g. an essay, painting, experimental design) is systematically rated on multiple criteria of quality. An analogy is the kind of scoring one sees in Olympic ice skating or diving by panels of experts.

The main virtue of performance testing can be seen in contrast to the limitations of current forms of multiple-choice testing. The multiple-choice format of existing mandated testing programs limits the forms of knowledge and skills that are measured mainly to recall and recognition of facts, simple concepts, and formulas, and to certain types of inferential and analogical reasoning skills. (Also, however, standardized tests measure such confounding factors as "test-wiseness," guessing, coaching, and unfortunately the ease of

cheating under multiple-choice answer formats.)[19] While research is steadily improving the quality of multiple-choice items and multiple-choice questions are not without value, the fundamental limitation is the student is not required to go beyond picking out from pre-given choices the best answer. Thus, large-scale, mandated, multiple-choice testing -- our main instrument for publicly certifying student competence and evaluating school quality -- virtually ignores a broad range of the higher order knowledge and abilities public schools are supposed to teach.

The main obstacle to large-scale performance testing is cost. Compared with multiple-choice tests, performance assessment is very expensive. Performance tasks take considerable time to develop and validate. Insuring reliable measurement for performance tasks is very labor intensive and problematic: if raters' standards of evaluation vary across students or sites, serious issues of fairness and equity are raised. If raters' standards shift to meet student performance, the potential for measuring conceptual understanding and applications may give way to assessing primarily factual recall and skills.

These problems are not insurmountable and are currently being solved in many innovative ways.[20] The most widely used form of performance assessment is

---

[19]R. Sarnacki, "An Examination of Test-wiseness in the cognitive domain." Review of Educational Research, vol. 2, 1979, pp. 252-279.

[20]Edward Roeber, "Performance Assessment: It's Real and Feasible -- But Now It Is Critical." Paper presented at the UCLA Quality Indicators Conference, October, 1989. Having used performance assessment for over a decade, Alverno College in Milwaukee is a nationally recognized leader. See, Russell Edgerton, "Abilities That Last a Lifetime: Alverno in Perspective." Bulletin of the American Association for Higher Education, vol. 36, No. 6, pp.

in writing, primarily because multiple-choice tests obviously cannot assess writing. Direct assessment of writing has a well-accepted technology and track record. It is currently a component of 28 state assessment programs. Performance assessment in other areas is growing. New York now assesses all 4th graders yearly on laboratory and "hands-on" science skills. Essay exams have been part of the Regents examination system since the turn of the century. California recently piloted open-ended, multi-step mathematics problems on its state test. Connecticut throughout the 80s has assessed foreign language speech; different forms of writing; and has developed methods to assess the performance of students in groups on collaborative problem solving exercises. Vermont is using student portfolios to assess engagement, progress, and performance in writing and mathematics. The federal government has already piloted and soon will be using performance-based tasks on the National Assessment of Education Progress.[21]

As with all forms of testing, "teaching-to-the-test" can be a potential problem for performance assessment. That a given performance task may be a more authentic demonstration of academic achievement is little consolation if it is produced by highly "targeted" instruction to the neglect of other important learning goals. However, if a performance assessment task constitutes a genuine intellectual challenge it should be "taught to" as Wiggins proposes in his definition of a "true" test.

---

1 - 6, 1984.

[21]Interestingly, the National Assessment of Educational Progress in pioneering work used constructed response items in its 1971 large scale assessment. Edward Roeber, Director of Assessment, Michigan State Department of Education, Ann Arbor, MI, personal communication, December, 1990.

> The true [performance] test is so central to instruction that it is known from the start and repeatedly taken <u>because</u> it is both central and complex -- equivalent to the game to be played or the musical piece to be performed. The true test of ability is to perform consistently well tasks whose criteria for success are known and valued (p. 706).[22]

**Indicator Systems**

The concept of "educational indicators" carries the quest for getting more information on school quality another logical step. If the analogy behind the standardized test is the thermometer, and behind performance assessment is the panel of expert raters, then, the analogy behind indicator systems is the "consumer reports" profile.[23]

The theory behind indicator systems is that, rather than relying exclusively on test scores, policy-makers and the public should have fuller portraits of school, district, and state educational performance. Indicator reports can include whatever facts and figures central authorities believe is of interest, but information related to educational quality usually is of most

---

[22]Grant Wiggins, "A True Test - Toward More Authentic and Equitable Assessment," <u>Phi Delta Kappan</u>, May, 1989, pp. 706-713.

[23]For more on indicator systems, see: Jeannie Oakes, <u>Educational Indicators: A Guide for Policymakers</u> (Center for Policy Research in Education, Rutgers University, 1986); Stephen Kaagan and Richard Coley, <u>State Education Indicators: Measured Strides and Missing Steps</u> (Center for Policy Research in Education, Rutgers University, and Policy Information Center, Educational Testing Service, Princeton, New Jersey, 1989).

interest: test scores, student retention rates, post-graduation outcomes, participation indicators (e.g., participation in the Advanced Placement program, National Merit Scholars competitions, community service activities), expulsion and suspension figures, measures of student attitudes and school-community relations.

Aggregated to the state level quality indicators provide a "condition of education" report. Such reports are available for the nation[24] and for most states, although the quality of state reports varies greatly both in their organization and clarity and in the usefulness of their information. For an indicator to be meaningful, among other things, it must be based on clear and consistent definitions (e.g., of "dropout") and reliable record-keeping.

"School Report Cards," an idea growing in popularity, refers to the development and publication of school level quality indicators (sometimes in booklet form) for use by the public and by policy-makers. Given the large volume of school, student, and staff information state agencies collect and their growing computerized information storage and processing capabilities, school report cards are yet another step in the continuing interest for information for accountability, program improvement, and other school reform purposes.

---

[24]National Center for Education Statistics, The Condition of Education: Elementary and Secondary Education (Washington, DC, 1990.)

## CURRICULAR AND ORGANIZATIONAL REFORMS: TECHNOLOGIES TO RAISE TEST SCORES

Declining test scores defined the crisis in education and fueled the education reforms of the late 70s and early 80s. Reformers expanded testing programs to implement consistent standards across schools, to evaluate the quality of schools, to guide planning for low-scoring programs or students, and to certify whether or not students were educated. It is only natural that some of the most popular curricular and organizational reforms were in their essence, technologies to raise test scores.

In these reforms the image of the test as a device of measurement begins to fade. In reforms involving principles of effective schools theory, mastery learning, and curriculum alignment, the test is less an external measure than an integral component of the reform.

### Effective Schools Practices

In the mid-70s, the "schools don't make a difference" conclusion -- an over-simplification of findings from ground-breaking studies lead by James Coleman, Christopher Jencks, and others -- was widely cited.[25] These researchers found that after accounting for family background variables, the school variables they measured (mostly school resource variables) had very little differential effect on achievement (usually measured by standardized test scores). Effective schools research probed this issue in more depth by

---

[25]James Coleman et. al. Equality of Educational Opportunity (Washington, DC: U.S. Government Printing Office, 1966); Christopher Jencks et. al. Inequality: A Reassessment of the Effect of Family (NY: Basic Books, 1972).

looking for associations among measures of organizational processes and characteristics and test scores under conditions where school student composition could be controlled. Effective schools research identified organizational and pedagogical practices characteristic of "positive outlier" schools -- schools with test scores exceeding their predicted test scores given the socio-economic status of their student body.

In short, researchers asked, what is different about schools with poor children that have unusually high test scores? In general, this is what they found:

- o strong instructional leadership on the part of the principal;
- o a clear and consensual school mission;
- o a collective commitment to high academic expectations and the belief that all students can succeed;
- o a safe and orderly school climate;
- o close monitoring of student achievement for program evaluation.

Effective schools theory which achieved extraordinary visibility in 1979,[26] helped set the tone for the aggressive top-down education reforms of the early 80s. "Effectiveness" meant goal-oriented management, a more uniform curriculum, test-based decision making, and above all, higher test scores.

---

[26]William, Brookover et. al., School Social Systems and Student Achievement: Schools Can Make a Difference, (NY: Praeger, 1979); Ronald, Edmonds, "Some Schools Work and More Can," Social Policy, vol. 9, 28-32, 1979; Rutter, Michael, et. al., Fifteen Thousand Hours: Secondary Schools and Their Effects on Children, (Cambridge, MA: Harvard University Press, 1979).

(However, the specific goal of boosting scores of low-achieving students prominent in effective schools theory was less of a force in the reforms of the 80s.) This reinforced acceptance of standardized achievement test scores as a legitimate measure of organizational performance.

That effective schools theory achieved such rapid and widespread acceptance attests to the powerful demand for a formula to boost test scores. One 1984 federal study reported 1,750 districts claiming effective schools programs, and Education Commission of the States estimated in 1986 about 20 states had policies promoting effective schools practices.[27] The diffusion of effective schools principles was helped by states which mandated formal (and documented) school improvement planning for schools with low test scores. New York's "Regents Action Plan," for instance, required that in "low performing schools" identified by the state's Comprehensive Assessment Report, "a self-improvement plan will be required. If sufficient progress has not been made in correcting the deficiencies, the Department will require corrective measures targeted at the specific deficiency."[28]

While the heyday of effective schools research is past, its influence on education policy continues and has been profound. One reason is that effectiveness theory gave board officials, administrators, and teachers alike a common, comprehensible, and credible set of principles to follow in efforts

---

[27]Olson, Lynn, "Effective Schools." Education Week, January 15, 1986, pp. 11-21.

[28]New York State Board of Regents, Proposed Action Plan To Improve Elementary and Secondary Education Results in New York (Albany, New York, 1984).

to make schools better. Regular school improvement planning and workshops, based heavily on school effectiveness research, has been incorporated into policy in most states.

Curriculum Alignment

As state authorities in the late 70s and early 80s became more involved in school reform and as higher test scores became increasingly the avowed goal of reform,[29] attention to curriculum alignment grew. Curriculum alignment requires a clear and specific, written curriculum -- a framework of topics and objectives to which textbooks, lessons, pedagogy, and tests are matched. The idea is straightforward. If the goal is to improve test scores, then instruction should focus on what is tested.

Mastery learning, with origins in the 1960s,[30] is the main progenitor of curriculum alignment reforms. Mastery learning is a highly scripted approach

---

[29]From the cover of a Florida policy document: "On a statewide average, educational achievement in the State of Florida will equal that of the upper quartile of states within five years as indicated by commonly accepted criteria of achievement" (i.e., test scores.). Student Performance Standards of Excellence in Mathematics, Science, Social Studies, and English (Tallahassee, FL: Department of Education, 1984). For an informative article on curriculum-testing alignment in a large Maryland school district, see Robert Rothman, "District Ties Goals to Test Scores," Education Week, March 22, 1989, p1.

[30]See, Benjamin Bloom, "Learning for Mastery, UCLA Evaluation Comment, 1, 1968, pp. 682-688; Carroll, J. B. "A Model for School Learning," Teachers College Record, 64, 1963, pp. 723-733; and, Keller, F. "Good-bye, teacher..." American Journal of Applied Behavioral Analysis, vol. 1, 1968, pp. 79-89.

to teaching, starting with objectives, specifying levels of performance that constitute mastery, delivering instruction aligned with the objectives, and monitoring student performance with tests also aligned to the objectives. A recent review of 108 studies of mastery learning selected for their methodological quality found "the average student in a mastery learning class performed at the 70th percentile, whereas the average student in a class taught without a mastery requirement performed at the 50th percentile" (p. 271).[31]

While there is an established tradition of curriculum alignment strategies at the classroom and school level, at the district and state level, it is a newer concept, many more variables are involved, and alignment is less easy to achieve. Not only have state education agencies had to contend with traditions of local curriculum autonomy, the local variables that influence course content and classroom instructional practice are not easily reached by state policies. Nonetheless the press for statewide education reform and higher test scores has been strong, and states have gradually tightened control over the curriculum variables they can influence. Districts under pressure to raise test scores on state tests have done the same.[32]

---

[31]Chen-Lin C. Kulik, James Kulik, and Robert Bangert-Drowns, "Effectiveness of Mastery Learning Programs: A Meta-Analysis," Review of Educational Research, vol. 60, No. 2, pp. 265-299.

[32]Ken Komoski cited by Lynn Olson, "Districts Turn to Nonprofit Group for Help in 'Realigning' Curricula To Parallel Tests." Education Week, Oct 28, 1987, p17,19. Textbook manufacturers market their books in "big-market" states and districts by demonstrating (in documentation and in sections of the books themselves) the alignment of their textbook content with state curriculum frameworks through "correlational analyses."

Policy-makers in the 80s pushed for "state-powered curriculum reform" and used the instruments at their disposal for curriculum control purposes.[33] Courses required for graduation were increasingly dictated by state policy.[34] State education agencies across the country rewrote -- or wrote for the first time -- curriculum frameworks, making them more prescriptive.[35] Many states with textbook adoption policies began scrutinizing more closely the match between their textbooks and their curriculum guidelines.[36] General "off-the-shelf" standardized achievement tests -- the staple of state testing for decades -- increasingly were augmented by custom-developed tests designed to assess state curriculum guidelines and goals.

---

[33]Larry Cuban, "State-Powered Curricular Reform, Measurement Driven Instruction," The National Forum, vol. LXVII, No. 3, 1987, pp. 22-25; see also: Michael Kirst, "Instructional Leadership at the State Level: What is the New Focus? NASSP Bulletin, vol. 71, No. 498, April, 1987, pp. 49-54.

[34]William Clune with Paula White and Janice Patterson, The Implementation and Effects of High School Graduation Requirements: First Steps Toward Curricular Reform (New Brunswick, NJ; Center for Policy Research in Education, Rutgers, The State University of New Jersey, 1989, ED 304 756).

[35]In a survey of 27 state social studies specialists, 26 said course requirements and guidelines had become more specific in last 4-5 years. The investigators concluded, "Despite great differences among the states, a very strong generalization emerges from the study, namely, that the current 'flavor' of social studies throughout most of the country is highly prescriptive. Many prescripts have been applied in recent years to students, teachers, and curricula." Council of State Social Studies Specialists, Social Studies Education, Kindergarten - Grade 12 (National Council for the Social Studies, Washington, DC, 1986).

[36]Harriet Tyson-Bernstein A Conspiracy of Good Intentions. Washington D.C.: Council for Basic Education, 1988; Harriet Tyson-Bernstein, "Three Portraits: Textbook Adoption Policy Changes in North Carolina, Texas and California." Occasional Paper, Institute for Educational Leadership, Washington, DC, 1989.

New York's Regents program, which has recently expanded its testing but is not a product of recent reforms, is a good example of comprehensive curriculum alignment at the state level. New York has lengthy and detailed curriculum guides and year-end examinations for all academic courses required for a Regents diploma. The guides are used as a basis for selecting textbooks, conducting inservices, and developing instructional materials;[37] and the Regents exams, are used as the basis of awarding Regents credits for graduation. (Regents exams are not required; about 60% of students take at least one exam, and about half of high school graduates get a Regents diploma.) California is another state that has made comprehensive curriculum alignment a priority. (Appendix A shows key elements of California's curriculum policy alignment system.[38] Appendix B shows selections from a Florida district's curriculum guide, which is part of a curriculum alignment system at the district level. The excerpts are from American History.)

In practice, curriculum alignment can range from state officials selecting a standardized norm-referenced test based on its match with loosely-defined state education goals to exhaustive content analyses assessing detailed matches among test, curriculum, and textbook objectives. The prevalence, extent of coordination, and effects of state and district curriculum alignment efforts is difficult to assess and there is little research here to draw on. However, 35 states test students using state-

---

[37]Doug Archbald, A Description and Analysis of the Purposes, Prescriptiveness, and Possible Effects of State Curriculum Guides in Mathematics and Social Studies (working title), (University of Wisconsin-Madison: Center for Policy Research in Education, forthcoming, 1991).

[38]Source: Michael Kirst, op. cit.

developed or state-selected tests and assess their performance against state-established performance standards.[39] So curriculum alignment exists to some degree in a majority of states.

## CONCLUDING COMMENTS ON THE ROLE AND INFLUENCE OF MANDATED TESTING IN EDUCATION REFORM

What has been the role and influence of mandated testing in education reform? The concluding portion of Part 1 addresses this broad topic by commenting first on the role of mandated testing in education reform, and then by presenting some evidence related to achievement outcomes.

### Mandated Testing in the Education Reform Process

Mandated testing and top-down education reform have had a strong symbiotic relationship for almost two decades. The impetus for reform has come consistently from outside the education system and has risen in part from diminished confidence of reform leaders in government, business, and universities in the ability of school people to initiate needed educational improvements. Scores on nationally standardized (though not necessarily mandated) tests have played a crucial role in mobilizing support for reform.

Reform sprung in the early 70s from reformers turning declining test scores into a basic skills and literacy crisis. This helped initiate the

---

[39]Margaret Goertz, op. cit.

minimum-competency testing movement. Reformers in the 80s also cited low test scores and raised the specter of a "rising tide of mediocrity," foreign economic domination, and inept workers unable to meet the demands of an increasingly "high-tech" world.

Reformers' and policy-makers' desire to impose external and publicly visible standards on local authorities has contributed to increased mandated testing. With the weight of state authority and the approval (or at least acquiescence) of the public, scores on mandated tests have acquired enormous weight in making final judgments about quality, performance, and progress -- often being the only standard used. State and district superintendents' public proclamations for all out efforts to raise test scores have won them political support and media approval. Districts, schools, and teachers producing steep gains in test scores have won national acclaim and awards and in some cases assumed almost legendary qualities. Writes the co-director of the Center for Research on Evaluation, Standards, and Student Testing, "current interest in test scores is at an all time high (p. 127)."[40]

Inevitably, as scores on mandated tests have acquired legitimacy and evaluative weight, mandated testing has become more of a policy lever. "Testing has changed dramatically from its former role as an index of educational progress... informing the public has taken a back seat to driving

---

[40]Robert Linn and Stephan Dunbar, "The Nation's Report Card Goes Home: Good News and Bad About Trends in Achievement," Phi Delta Kappan, vol. 72, No. 2, October, 1990, pp. 127-133.

policy and influencing practice (p. 769, 770)[41]." Policy-makers have attached rewards and sanctions to performance on mandated tests on the assumption that this will motivate students to achieve and encourage teachers to teach state prescribed curriculum goals and basic skills objectives.

That existing mandated tests neither measure nor disclose very well what students and schools accomplish in problem-solving and conceptual understanding has unfortunately received comparatively little attention in reform policies. Research and development work in the area of performance testing and in educational indicator systems reflects growing recognition of the limitations of existing uses of mandated testing.

**Mandated Testing and Achievement Goals of Reform**

It is tempting to attempt to identify a discrete, separable influence of tests in the reforms discussed thus far and to render a judgment about whether this influence helps or hurts the reforms. This would not only be premature, but overly simplistic. Testing is part of a whole constellation of forces shaping education in our schools. The nature and extent of mandated testing is very much a result of widely held beliefs about curriculum, teaching, and learning. We return to this notion later in this report. At this point, we adopt the standards of the top-down reforms and suggest these reforms -- with mandated testing contributing -- have had some noteworthy successes. But, like many, we believe limitations of existing mandated testing and the top-down

---

[41]Mary McClellan, "Testing and Reform," Phi Delta Kappan, vol.69, No. 10, June, 1988, pp. 768-771.

model make it unlikely that continuing in this direction will prove fruitful given the educational needs facing our country.

Evidence from test scores suggests curriculum-testing alignment raises test scores. A Florida district where we are currently conducting research instituted a curriculum alignment system in the late 70s. It uses prescriptive course-level curriculum guides, which highlight state and district curriculum objectives and identify for the teacher where these objectives are covered in the course textbook (Appendix B). Textbooks are selected by district committees, with a single textbook adopted for each course. The district has course-end "minimum skills tests" which students must pass to receive credit for each course they take. The district's test for its lower level courses are similar in content to Florida's state test, the SSAT. The SSAT includes a functional literacy portion all students must pass to earn a diploma. The district also administers the Stanford Achievement Test in grades 1 - 12.

The district's test scores between 1977 and 1987 went up steadily in every category. On the SSAT, the percent of students passing state standards in five different grades in mathematic rose by an average of 14 percentage points; in communications, the average increase was 13 points over the 10 years. Score increases on the Stanford Achievement Test across 12 grades in mathematics and reading ranged from 7 points to 24 points. Test score gains of equivalent magnitude were registered on the district's course-end "minimum skills tests."

Although this Florida district may not be representative of all

districts that pursue curriculum alignment, this district is by no means an anomaly. Other districts and states have posted similar gains. Scores on Florida's SSAT have gone up by as much as 30% (in mathematics) since the test's inception. Scores on the Texas Assessment of Basic Skills in mathematics rose from 70% of 9th graders achieving mastery in 1980 to 84% in 1985. On the reading portion, scores increased from 70% to 78%.[42] Table 4 shows score gains in other places.[43] It is very likely these kinds of test score gains have occurred more widely in districts and states where curriculum-testing alignment has been systematically pursued.

[ TABLE 4 HERE ]

Research has not shown precisely how this is accomplished and whether these kinds of test score gains indicate that students have actually learned more worthwhile knowledge and skills. These issues are discussed in Part 2.

These magnitudes of test score gains have certainly not occurred in all districts or states. However, there is some evidence that collectively the kinds of standards-raising and curriculum control reforms we have discussed have contributed to some achievement gains sufficient to register on national tests. Trends in achievement in the nationally representative samples of students tested by National Assessment of Education Progress and the Iowa Test

---

[42]Office of Technology Assessment, op. cit., p. 272.

[43]Source for Table 4: W. James Popham, "The Merits of Measurement-Driven Instruction," Phi Delta Kappan, vol. 68,. No. 9, May, 1987, pp. 679-682. Note, ranges in rightmost column, denote the range of percentage increases across the different grade levels and tests in columns on left.

**TABLE 4 Improvements in Student Achievement Associated With Curriculum Alignment**

| Locale | Subj. | Grade | Period | Gain(%)* |
|---|---|---|---|---|
| Ala. | 3R's | 3,6,9 | 1981-86 | 1-13 |
| | 3R's | 11 | 1983-85 | 4-8 |
| Conn. | 3R's | 9 | 1980-84 | 6-16 |
| Detroit | 3R's | 12 | 1981-86 | 19 |
| Md. | 3R's | 9 | 1980-86 | 13-25 |
| | Soc. St. | 9 | 1983-86 | 23 |
| N.J. | Rdg/Math | 9 | 1977-85 | 16-19 |
| | Rdg/Math | 10 | 1982-85 | 8-11 |
| S.C. | Readiness | 1 | 1979-85 | 14 |
| | Rdg/Math | 1-3,6,8 | 1981-86 | 12-20 |

*Figures represent the increased percentage of students who have mastered standards of quality during the period in question.

of Basic Skills show a general pattern of increasing scores from the mid-70s to the late 80s.[44] The increases are not dramatic, but they are not insignificant either. That the increases are greatest in the basic skills areas and among minority students, might be accounted for in part by the generally greater tightening of state and district control over minimum achievement standards. Basic skills instruction is the easiest to monitor with mandated testing and control with curriculum and accountability policies.

The reason these test score gains are encouraging is that national tests are not curriculum-aligned -- that is, while national tests reflect common curriculum goals which schools in general pursue at particular grade levels, these tests are not designed to reflect any particular district's or state's curriculum. Thus, while it is possible for a particular district to follow closely a prescribed curriculum which is assessed by an aligned test, schools and districts across the country are not similarly oriented toward any particular national curriculum. National tests theoretically sample content domains that are less susceptible to coaching and short-term learning.

## A Long Way From Excellence

Unhappily, in spite of whatever modest achievement gains may have occurred since the mid-70s, when achievement is looked at in relation to actual performance criteria or in relation to the test achievement of other countries, there is still reason for serious concern about low quality

---

[44]Robert Linn and Stephan Dunbar, op. cit.

education. High school students from the United States continue to fare poorly, near the bottom, in international comparisons of achievement. High school seniors from West Germany or Japan would compete well with American students in their first or second year of college.

The 1988 National Assessment test found, for instance, only 6% of 17 year-olds could solve the following multi-step math problem.

> Christine borrowed $850 for one year from the Friendly Loan Company. If she paid 12% simple interest on the loan, what was the total amount she repaid?

Only about 5% of 17 year-olds could interpret a paragraph that begins, "In the years between 1940 and 1960, literature, the arts, and culture in general became increasingly oriented to the many. In an economy of high productivity, deluging millions of people daily with movies, magazines, books, and television programs, American culture achieved a degree of homogeneity never dreamed of before."[45]

There is little evidence that what the National Commission on Excellence in Education wrote in 1983 (p. 9)[46] would not be just as true today.

---

[45]These examples are from Arthur Applebee, Judith Langer, and Ina Mullis, Crossroads in American Education: A Summary of Findings (Princeton, NJ: National Assessment of Educational Progress and Educational Testing Service, 1989, p.29, 51)

[46]National Commission on Excellence in Education A Nation At Risk: The Imperative of Educational Reform. Washington, D.C.: U.S. Government Printing Office. April, 1983.

> Many of the 17-year-olds do not possess the "higher-order" intellectual skills we should expect of them. Nearly 40 percent cannot draw inferences from written material; only one-fifth can write a persuasive essay; and only one-third can solve a mathematics problem requiring several steps.

Top-down reforms have expanded mandated educational testing and strengthened central control over curriculum. Increased mandated testing appears to have helped improve accountability for basic skills and helped promote basic skills remediation for low-achieving students. Many states and districts have reported increased test scores. However, existing mandated testing programs are based almost entirely on multiple-choice testing which do not effectively measure the higher order knowledge and skills schools are supposed to teach. National and international tests with more challenging content suggest we have a long way to go in education reform.

## PART 2:
## CRITICISMS OF STANDARDIZED TESTS AND GROWING INTEREST IN RESTRUCTURING

### INTRODUCTION

Since the mid-70s testing, as discussed above, has been an integral fixture in centralized, top-down reform. In part from reformers' disenchantment with the apparently meager test results from the top-down

reforms and in part from growing political influence of coalitions of testing opponents, the late 80s brought renewed and increasingly strong criticism of mandated testing and calls for a more radical "restructuring" of public education. This section will discuss two main criticisms of testing, suggest the empirical support for these criticisms is far from solid, and describe several important restructuring reforms that fall outside of the top-down model.

## CONCERNS ABOUT ADVERSE EFFECTS OF MANDATED MULTIPLE CHOICE TESTING

Educational testing has been the subject of controversy in one form or another for decades. However, the recent eruption is one of the louder and has spread from the staid confines of scholarly conferences and journals to front pages and the vehicles of heavyweight reform leaders. The Chairman of the National Commission on Testing and Public Policy just declared, "...the testing enterprise has in many instances gone haywire and is driving our educational system in the wrong direction."[47] A report from the National Center on Fair and Open Testing contends, "standardized tests undermine school improvement instead of advancing its cause."[48] The National Council's for each of the academic subjects, the presidents of the Carnegie Foundation for the Advancement of Teaching, the National Education Association, and the American

---

[47]Barnard Gifford, cited by Jean Evangelauf, in "Reliance on Multiple-Choice Tests Said to Harm Minorities and Hinder Reform; Panel Seeks a New Regulatory Agency." Higher Education Chronicle, Vol. 36, No. 37, May 30, 1990, p1, 31.

[48]Monty Neill and Noe Medina, "Standardized Testing: Harmful to Educational Health," Phi Delta Kappan, vol. 70, No. 9, May, 1989, pp. 688-697.

Federation of Teachers, and prominent leaders from many other national organizations of educators have called for roll backs and reforms in state mandated testing programs.[49] (Performance testing and indicator systems are not the target of anti-testing sentiment.) Doubtless, the escalation of mandated testing in the 80s and the current eruption are linked.

Two main lines of argument are advanced in criticisms of mandated educational testing.[50] The first is that under certain conditions testing adversely affects curriculum and instruction. The argument is as follows: Mandated student testing is conducted almost exclusively using facts and skills-dominated multiple-choice tests. Because there is accountability pressure for schools to achieve high test scores (or at least avoid low scores), teachers are forced to "teach to the test" -- that is, to shape their curriculum and instruction around the goal of developing students test-taking abilities. This means teaching in a way that emphasizes recall and recognition of facts and concepts, certain forms of verbal logic and inferential skills, and mathematical computational skills. What the tests do not directly measure does not get taught: creativity, depth of understanding, integration of knowledge, ill-structured problem solving, and communication. The end result is a curriculum consisting largely of basic skills instruction and unrelated

---

[49]"Coalition Calls for Genuine Accountability," Fairtest Examiner, vol. 4, No. 1, Winter, 1989-90, p1.

[50]A number of more specific, and often quite technical, criticisms have been made against testing. These criticisms are often specific to particular types of tests, or to particular uses of test results. These criticisms have to do with assumptions underlying the psychometric model, scaling procedures, predictive validity, and distinctions between ability and achievement. Addressing these criticisms are beyond the scope of this report.

fragments of information to be memorized for later recall on tests.

A second and related issue implicates mandated testing in the erosion of teacher professionalism. According to this critique, teachers' level of control over their work in their classroom has diminished as control has shifted to prescriptive testing, curriculum, and accountability policies. The use of tests to make evaluation decisions (e.g., promotion and retention of students, judgments about teacher or school quality) gives them clout, and "when such high stakes are attached to scores, tests can be expected to exert a strong influence on 'what is taught, how it is taught, what pupils study, how they study, and what they learn'."[51] This, in turn, undermines teachers' pedagogical autonomy and feelings of professional efficacy.[52] McNeil (1988:335) contends:

> By prescribing curriculum and instruments of assessment, such reforms...separate the craft of teaching from teaching style and remove teachers' discretion from their judgments about students and what they need to know. In this de-skilled model of teaching, one teacher lamented, the teacher becomes little more than an

---

[51]Linda Darling-Hammond, "Achieving Our Goals: Superficial or Structural Reforms?" Phi Delta Kappan, vol. 72, No. 4, Dec., 1990, pp. 286-295. (The embedded quote is from George Madaus et al., "The Sensitivity of Measures of School Effectiveness," Harvard Educational Review, vol. 49, 1979, pp. 207-230.)

[52]See also: Jack Frymier, Bureaucracy and the Neutering of Teachers, Phi Delta Kappan, vol. 69, No. 1, 1987, pp. 9-14; Linda Darling-Hammond, "Accountability for Professional Practice," Teachers College Record, vol. 91, No. 1, Fall, 1989, pp. 59-80.

assembly-line worker, performing mechanical tasks.

In this view, then, the growth of mandated testing and its use for accountability has been counter-productive: it adversely affects curriculum and the quality and status of working conditions for teachers. Critics contend test score gains mean little because they do not signify important forms of achievement, and central policy-makers' interest in accountability does not warrant the more important sacrifice of teacher autonomy.

## IS MANDATED TESTING DETRIMENTAL TO CURRICULUM AND TEACHERS?

The issue here is not, "Do mandated tests influence curriculum and instruction?" Clearly, as stated earlier, they do. The key questions are -- How much do tests influence curriculum? In what way? Are effects on curriculum necessarily bad? And how do teachers feel about preparing students for mandated tests?

The critical literature on mandated testing is primarily conjectural, much is based on anecdotal evidence. What research there is does not provide satisfactory answers and appears to paint more of a picture of weak than strong effects. Also, it is not clear that the mere finding of "effects" justifies the conclusion that the quality of instruction is worse where these effects occur.

**Effects on Curriculum**

One recent survey on the effects of mandated testing revealed that in nationally representative sample of 8th grade mathematics teachers (n = 552), almost half reported not preparing students at all for mandated tests, and of those that did report test preparation, almost half spent no more than several periods a year prior to test administration (and mathematics is one of the most tested areas).[53] Another study indicates some effects of testing, but does not shed light on the magnitude. It is based on surveys and interviews with 285 mostly elementary teachers in seven midwestern states and found,

> ...most teachers find standardized testing more helpful than hurtful. "Current testing," the teachers claimed, "orients teaching to stated goals and increases emphasis on basic skills. Only a few teachers saw such concentration no longer in the best interests of children. Also, the majority saw district-wide testing to be a curb on deviation from consensus goals.[54]

When one queries not teachers in general (i.e., random samples), but focuses on teachers in particular "high-stakes" testing conditions -- such as minimum-competency tests, school evaluation tests, or externally developed

---

[53]Thomas Romberg, Anne Zarrinia, and Steven Williams, The Influence of Mandated Testing on Mathematics Instruction: Grade 8 Teachers' Perceptions (National Center for Research in Mathematical Science Education, University of Wisconsin-Madison, 1989).

[54]Robert Stake and Paul Theobold, "Teachers in 7 States: 'Testing is Helping Schools.'" Educational Assessment and Policy Use (Newsletter of ECAP Researchers), no. 5, November, 1988.

course-end tests -- then testing exerts a greater influence on curriculum and instruction. A close-up study of four elementary classrooms with both mandated stating and district (objectives-based) testing found students spending up to 18 hours taking tests and about 54 hours receiving instruction that appeared to be directly oriented toward the tests over the course of one year.[55] New York's Regents courses (in most academic subjects) have state-developed course-end examination.[56] Grades on Regents exams determine whether students do or do not receive credits toward the state's Board of Regents diploma; passing rates for students also reflect on their teachers, so the stakes are high. Even under these conditions, teachers spend only several class periods to as much as about ten class periods reviewing and preparing for the examinations. Even the upper number as a percentage of total class sessions (around 175) reflects a rather modest direct effect of testing.[57]

Wilson and Corbett[58] conducted a study of mandated minimum competency testing in Maryland and Pennsylvania. Pennsylvania's test is used for identifying students needing remedial instruction in basic skills, and

---

[55]Claire Rottenberg and Mary Lee Smith, "Unintended Effects of "External Testing in Elementary Schools." Paper presented at the Annual Meeting, American Educational Research Association, Boston, MA, April, 1990.

[56]These exams, developed largely by teachers, consist of a variety of question types: multiple-choice, short answer, and essay. Quantitative problems in mathematics and science require students to show their work.

[57]Douglas Archbald, "Curriculum Control and Teacher Autonomy." Paper presented at American Educational Research Association annual conference, Boston, Massachusetts, April, 1990.

[58]Bruce Wilson and H. Dickson Corbett, Two State Minimum Competency Testing Programs and Their Effects on Curriculum and Instruction (Philadelphia, PA: Research for Better Schools, 1989).

districts' scores are publicly reported. Passing Maryland's test is a graduation requirement, indicating higher stakes conditions for the Maryland test. The survey found effects in both states, but stronger ones in Maryland where majorities of respondents reported using practice tests and reviewing course content and objectives prior to tests.

One of the larger studies to date on determinants of curriculum and instructional practice in mathematics used surveys, interviews, and teacher logs and involved 105 upper-elementary teachers. It concluded, "Little evidence exists to support the supposition that national norm-referenced, standardized tests administered once a year have any important influence on teachers' content decisions (p. 9)"[59] The authors note that "curriculum-embedded" tests have important effects, but only "when they have been explicitly tied to the curriculum and when they are readily accessible and easily used by teachers." The authors stressed that in relation to teachers' own discretion to decide content and methods, tests had little explanatory power in accounting for what was taught and how it was taught.

These studies, though the findings are somewhat disparate, indicate under high-stakes conditions mandated testing can influence content and instruction, but it is unclear how good or bad this influence is. Given materials, a set of clear content objectives, "stakes," and a mandated test, teachers will teach accordingly, trying to help students pass or score well.

---

[59]Andrew Porter, Robert Floden, Donald Freeman, William Schmidt, and John Schwille, "Content Determinants in Elementary School Mathematics." In D. A. Grouws & T. J. Cooney (eds.), Perspectives on Research on Effective Mathematics Teaching (Hillsdale, NJ: Erlbaum, 1988, pp. 96-113).

But even under the high-stakes conditions the influence of mandated tests still may be a small percentage of total instructional time. Effects are confined to the specific grades and content or skills tested and to a relatively short period of time prior to the test. In addition, it is preponderantly only low-achieving students experiencing test preparation, often in the form of remedial classes. While some view critically the very existence of remedial classes, such classes existed well before the spread of mandated testing.

The more important (and more difficult) question is -- is test preparation necessarily inferior instruction? Critiques of mandated testing commonly charge tests with "corrupting" instruction.[60] Critics argue or imply mandated testing prevents good teaching, making it mechanical, superficial, and fragmented. According to this view, but for mandated testing, curriculum and instruction would engage students more in problem-solving, creative activity, and analytical thinking projects.

The problem this poses for research can be phrased this way: would observers discover inferior curriculum and instruction in classrooms where mandated testing occurs compared with classrooms without mandated testing. Mandated testing under the right conditions can induce teachers to rearrange topics, to pay more attention to some skills or concepts or some students than others, and to undertake year-end reviews of subject matter; but it requires a significant inferential leap to assume mandated testing substantially worsens

---

[60]But see, Oran Stewart and Dan Green, "Test-taking Skills for Standardized Tests of Reading." Reading Teacher, March, 1983, pp. 634-638.

curriculum and pedagogy.

Mandated testing, especially when it emphasizes basic skills, is consistent with standard instructional practice in schools. It has not been used as a force for significant change (beyond the modest impacts described in different sections of this report.) In fact, research has consistently noted that when tests are proposed that genuinely disturb the status quo in schools, the result is that the test far more than standard practice gets changed. "Even the most orthodox and doctrinaire justification of cut-scores in terms of skills and competence is moderated in the end by consideration of pass-fail rates. Norm-referencing drives out criterion-referencing. Pure criterion-referencing exists only in textbooks and scholarly journals; it is not found in the world of practice."[61] Reformers who have advocated the use of "exhibitions of mastery" -- performance tests of literacy and academic competence -- for graduation from high school have observed the great difficulty this provocative idea has had in gaining acceptance and in being implemented in a manner true to its principles.[62]

Existing mandated testing programs, then, at best may contribute to some extra basic skills achievement in school systems. They probably in some classes shift resources and attention during testing periods to low-performing students and make it somewhat less likely that students with severely deficient skills "slip" through the system. Conversely, these tests, then, may

---

[61]Glass and Ellwein, op cit., p. 4.

[62]Grant Wiggins, "Thoughts on Obstacles to Curricular Reform," unpublished manuscript, 1986.

shift attention away under certain conditions from the needs of higher-performing students. For the teacher this is disruptive, and for the average or bright student this may create more unguided time or more boring drill and practice work. We are skeptical mandated testing seriously affects conditions beyond this, such as science instruction, literature, a given teacher's propensity (or lack thereof) to analyze issues in social studies, or the likelihood that a teacher will use cooperative learning strategies, a project-oriented curriculum, or other innovative pedagogical approaches. If teachers want to do these things they will, with or without mandated testing.

## Effects on Teachers

The other issue in mandated testing concerns effects on teachers and how they view their work. Evidence is equivocal. Some qualitative research has elicited strong complaints by teachers about mandated testing programs, or at least "substantial ambivalence."[63] The study (described above) of elementary teachers in two schools with high-stakes testing suggested teachers were under great pressure to produce high scores and were generally feeling disillusioned and skeptical about the value of test-oriented instruction.[64] On the other hand, some surveys find mixed or supportive attitudes about mandated testing.

---

[63]See, Susan Rosenholz, "Education Reform Strategies: Will They Increase Teacher Commitment?" American Journal of Education, August, 1987, pp. 534-562; and, OERI State Accountability Group, Creating Responsible and Responsive Accountability Systems. (Washington, DC: US Department of Education, 1988, p. 33.).

[64]Mary Lee Smith, "Meanings of Test Preparation," Paper presented at the Annual Meeting, American Educational Research Association, Boston, MA, April, 1990.

In a survey of a random sample (n=3300) of K - 12 teachers, only 19% rated as "too great" the "amount of [standardized] testing in your system;" 69% reported "about right" and 7% reported "too little." On the hypothetical question, "Do you personally favor a moratorium on...." 16% reported favoring a moratorium on "all standardized test," and 31%, on "state-mandated achievement tests."[65]

A current study by the Center for Policy Research in Education compares the curriculum autonomy and perceptions of high school teachers, using both interviews and surveys, under centralized and non-centralized curriculum control conditions. The amount of mandated testing and the use of test results is a major distinguishing feature between the centralized and the non-centralized conditions. Survey scales measuring "sense of efficacy" and "job morale" produced very few meaningful differences between centralized and non-centralized conditions.[66] Teachers working under conditions of central control in interviews and surveys described themselves as having almost complete control over pedagogy,[67] just as did the teachers who worked in districts

---

[65]Fran. Stetz and Michael Beck, "Attitudes Toward Standardized Tests: Students, Teachers, and Measurement Specialists." Measurement in Education, Vol 12, no. 1, pp. 1-11, 1981. Other surveys have found similar results. See, John Ward, Teachers and Testing: A Survey of Knowledge and Attitudes. Report of the research department of the American Federation of Teachers, July, 1980; and Stake and Theobold, op. cit.

[66]Douglas Archbald, 1990, op. cit.

[67]On a scale from 1 ("no control") to 6 ("complete control") over "setting standards of achievement in my class," teachers in the our two noncentralized districts rated themselves at 5.1 and 5.0, respectively, and in our two centralized districts at, 5.2 and 4.6, respectively. On the item "selecting teaching techniques," the noncentralized districts' means were 5.6 and 5.5, and the centralized districts' means were, 5.5 and 5.2.

without central curriculum control policies.

These studies highlight the need for additional research on intervening variables conditioning teachers' views of and responses to mandated testing. For instance, significant for teachers' feelings of efficacy and commitment is their prior experience with mandated tests. The Rosenholz study queried teachers in Tennessee during the first year of the new test. A certain amount of dissatisfaction and grumbling over new paperwork, changed routines, and imposed standards can be expected. Teachers in New York are used to Regents exams and the year-end regimen of preparation, grading, and paperwork they require. Thus, New York teachers relative to Tennessee teachers may work under more intense testing pressures and have a greater responsibility to prepare students, but the long tradition of state testing in New York creates expectations which make testing a taken for granted part of the job. Other variables conditioning teachers views and uses of mandated tests are also important, but have received little study -- like the quality of externally mandated tests, the quality of relations between state and local agencies, the level of state and local support for test-driven curriculum and remediation, and teachers' knowledge of standardized testing.

In sum, we suggest the case against mandated testing has been overstated. Mandated testing programs, while influential in some ways, do not have the pervasive curriculum and teacher control effects sometimes attributed to them. Critics have tended not to view the influence of mandated testing in the context of powerful forces like textbooks and curriculum mandates, teachers' own beliefs, talents, and interests, and larger organizational and

institutional factors beyond the reach of conventional policy.[68] Overall, despite the prodigious scale of the 80s top-down reforms and the steady growth since the 70s of mandated testing, teachers' control over what goes on in the classroom remains substantial, and curriculum and instruction remains largely unchanged. The 80s reforms have by and large left the core operations of teaching untouched.

## RESTRUCTURING: REFORMS THAT DECENTRALIZE CONTROL AND ACCOUNTABILITY

Mandated tests are easy targets for criticism and controversy: they are visible policy instruments, arbiters of opportunity, and they symbolize remote, technocratic central control. Even if mandated testing's actual influence on curriculum and instruction is weak, its real and symbolic role in growing top-down control is significant. This perception (reinforced by the lack of clear evidence of improvement in schooling) has spurred support not only for testing reform, but for a variety of reforms that share the goal of decentralizing authority and control and implementing structural changes in schools. School-based management and school choice, described next, are the two most prevalent reforms that seek to decentralize and restructure school governance and organization. Mandated testing as it is currently administered and used has little direct significance for these reforms.

---

[68]For an illuminating treatment of this, see, David Cohen, "Teacher Practice: Plus Ca Change," Michigan State University, East Lansing, The National Center for Research on Teacher Education, September, 1988, Issue Paper 88-3.

**School-based management**

School-based management refers to the delegation of decision-making authority to the school-site.[69] It is premised on two assumptions: (1) that those closest to the teaching and learning process can make the most effective educational decisions, and (2) that teachers' and principals' sense of efficacy and commitment to their work are likely to be greater if they have the authority and the discretion to make key educational decisions.[70]

> Jobs that give people autonomy and discretion require that they exercise judgment and choice; in doing so, they become aware of themselves as causal agents in their own performance. Loss of the capacity to control the terms of work or to determine what work is to be done, how the work is to be done, or what its aim is to be, widens the gap between the knowledge of one's unique contributions to work and any performance efficacy that can be derived from it."[71]

---

[69]School-based management's precursors include site-based budgeting and the decentralization initiatives of the 60s and 70s. The former was motivated in part by concerns over administrative efficiency, and linked to strategies like "management by objectives" and "zero-based budgeting;" the latter was tied in with the "community control" movement of the 60s, most notably in New York City.

[70]Studies have found positive and statistically significant relationships between measures of teacher efficacy and student achievement. For a discussion of school-based management theory see David, Jane, "Synthesis of Research on School-Based Management." Phi Delta Kappan, Vol 46, No. 8, May, 1989, pp.45-53.

[71]Rosenholz, op. cit., p. 540.

Reliable figures on the prevalence of school-based management are scarce. School-based management initiatives often take place "invisibly," on an individual school basis, and the concept itself is not easily defined. At least 100 districts nation-wide use school-based management. Among the better known are programs in Dade County, Tulsa Salt Lake City, Montgomery County (MD), and Cerritos (CA). States promoting school-based management through policy include California, Florida, Minnesota, South California, and Kentucky.[72] The scope of school-based management initiatives started in the Chicago schools in 1989 is truly unprecedented in decentralization reforms.

School Choice

Where site-based management is essentially a "democratic participation" theory of school-level organizational improvement, school choice is a market theory and assumes market mechanisms should play a stronger role in school improvement. According to this view, family choice, competition, and a reduced regulatory burden would improve public schools.

Magnet schools are by far the most common approach to school choice,[73] but recently other forms of school choice have emerged: 6 states have enacted

---

[72]"School-Based Improvement and Effective Schools: A Perfect Match for Bottom-Up Reform" An ACCESS Printout, Clearinghouse for Information About the Public Schools, The National Committee for Citizens in Education, Columbia, MD, 1988,.

[73]For figures on the prevalence of magnets schools and literature on their purposes and effects, see Douglas Archbald, "Magnet Schools: Equity, Choice, and Achievement Outcomes." Prepared for American Institutes for Research, Palo Alto, California, 1991.

and another 15 are considering inter-district choice options for all K-12 students; 17 states offer post-secondary enrollment options through which high school students can take college courses for credit at the state's expense.[74] There have been many proposals to give middle- and low-income families access to private schools through tuition subsidy policies, but so far only one has made it into law -- a 1990 law authorizing state tuition payments for about 1,000 low-income children in Milwaukee, Wisconsin to enroll in schools of their choice, public or non-sectarian private.

**Other Decentralizing/Restructuring Reforms**

"Deregulation" refers broadly to efforts to waive or eliminate bureaucratic regulations on the assumption that they can suppress or restrict local organizational or curriculum innovations. In theory, a school or district could be released from mandated student testing on the presumption testing prevents exploration of promising innovations. However, in practice, few districts have sought regulatory waivers for innovation.[75]

More than 20 states now offer schools and districts waivers from state regulations in areas such as staffing, curriculum, and budgeting.[76] To obtain a waiver, state officials typically require documentation of a school's

---

[74]National Governor's Association, Results in Education (Washington, D.C., 1989).

[75]Susan Fuhrman, "Diversity Amidst Standardization: State Differential Treatment of Districts." Paper presented at the Conference on Choice and Control in American Education, University of Wisconsin-Madison, May, 1989.

[76]National Governor's Association, op. cit.

capacity and commitment to implementing its plan. For instance, at the request of a school-site planning team, the Colorado State Board of Education granted a Boulder elementary school a waiver from the requirement that a principal hold an administrative certificate and that teachers be evaluated by an administrator. The school was then run by a three-teacher committee. The waiver is based on a 1989 state law allowing the board to waive statutes that "hinder or thwart educational improvement."[77]

## MINIMAL RELATIONSHIPS BETWEEN TESTING AND RESTRUCTURING REFORMS

Restructuring reforms bear little direct relationship to mandated testing. Testing is not a programmatic component of these reforms and a possible role for tests -- monitoring school-based management effects on test performance -- has not materialized. Certainly one reason for this is that a focus on short-term test score changes would detract from the more immediate organizational goals of decentralization of authority and control and empowerment of school level practitioners. However, even though the philosophy of school-based management is averse to the notion of evaluation by standardized test scores, it is likely school-based management will ultimately need to show results on quantifiable measures of student learning.[78]

---

[77]"Teachers to run Colorado." August 6th issue of EDCAL, newsletter of the Association of California School Administrators.

[78]In four literature reviews only two empirical studies of achievement outcomes are discussed (described in Malen, 1989). Neither study found achievement effects of school based management (using test scores as the dependent variable). Betty Malen, Rodney Ogawa, Jennifer Kranz, "What Do We Know About School-Based Management? A Case Study of the Literature -- A Call for Research." Paper prepared for the Conference on Choice and Control in American Education, University of Wisconsin-Madison, 1989. See, also, Jane David, "Synthesis of Research on School-Based Management," Education

Proponents of school choice believe information systems should make school quality indicators accessible and useful to parents to promote competitive incentives to improve schools. While some have advocated better and more accessible school-level performance indicators (advocacy which could dovetail with current interest and research on indicator systems, see Part 1), in practice this has not occurred.[79] In systems with magnet schools, however, school-level test information reportedly influences (among many criteria) parent decisions.[80] However, the role of test scores in school choice programs has never been closely studied.

## PART 3:

## THE NEED FOR A NEW VISION IN CURRICULUM AND TESTING

### INTRODUCTION

As we have seen, the last 15 years have been an active period of education reform. The first half of the 80s, building on the trends set by the minimum-competency testing movement, was dominated by top-down curriculum

Leadership, vol. 46, No. 8, May, 1989, pp. 45-53; and Lawrence C. Pierce, School Based Management, University of Oregon, Oregon School Study Council, 1980.

[79]Policy-makers have discussed developing a "consumer information" system (including school test score performance) to assist Minnesota parents' decisions in Minnesota's open enrollment program.

[80]Douglas Archbald, "Who Chooses Magnet Schools and Why?" Paper presented at the American Educational Research Association annual meeting, San Francisco, California, March, 1989.

control reforms which expanded mandated testing. They were put in place to increase standards, but they did not directly challenge the 1970s commitment to basic skills. As states increased requirements for students and schools, local discretion decreased. By the mid-80s, restructuring began to catch on. School-based management, teacher empowerment, school choice, and deregulation were touted as mechanisms needed to improve the quality of schools and of teaching and learning. In most cases, the earlier curriculum control initiatives remained in place. Some states opened up the possibility of waiving requirements for schools or districts who provided a convincing rationale; a few states waived requirements for schools demonstrating above average performance (e.g., South Carolina).

On the matter of curriculum, restructuring is largely silent. The assumption is that empowered teachers working in restructured schools will know what is best for students and do it. Others outside the restructuring movement have questioned this assumption, noting in particular the lack of clear goals of many restructuring initiatives.[81]

## A CONCURRENT CURRICULUM REFORM

---

[81]See for instance, Lynn Olson, "The Restructuring Puzzle," Education Week, Nov. 2, 1988, p.7. On school-based management, White writes, "Researchers, practitioners, and policy-makers interpret school-based management differently, and there are numerous variations within districts and schools regarding the levels of authority, the actors involved, and the areas of control." Paula White, "An Overview of School-Based Management: What Does the Research Say?" NASSP Bulletin, September, 1989.

In addition to the early 80s top-down standard setting reforms and the mid-80s restructuring reforms, there has also been the beginnings of an ambitious new curriculum reform. By the end of the 80s, professional societies, blue ribbon panels, and, at least one state (i.e., California) began calling for a fundamentally different curriculum from that typical of today's schools. Elsewhere we characterized the new goal as "hard content for all students."[82] To understand the significance of this new goal, its two parts must be considered in combination. By the first part, hard content, we mean conceptual understanding and application, not just facts and basic skills. Hard content emphasizes students learning to produce knowledge, rather than simply reproduce knowledge, to reason and to be analytical. The second part of the goal is that all students are to have the opportunity to learn hard content, not just academically elite students.

The beginnings of this new curriculum reform can be seen in a number of sources and for a number of academic subjects. The most visible evidence is for mathematics, where three publications, all appearing in 1989, call for an increased emphasis upon problem solving and application and a decrease in emphasis upon computational skills: Everybody Counts, published by the prestigious National Research Council; Science for All Americans, published by the American Association for the Advancement of Science; Curriculum and Evaluation Standards for School Mathematics, published by the National Council

[82]Andrew Porter, Douglas Archbald, & Alex Tyree, "Reforming the curriculum: Will empowerment policies replace control?" In S. Fuhrman (ed.)., The Politics of Curriculum and Testing (London: Taylor & Francis Ltd, in press).

of Teachers of Mathematics.[83] Science for All Americans advoactes ambitious content reforms in science for all students. Social studies experts are more concerned with arguing about the appropriate emphasis to place on history, but they, too, published three reports in 1989, and which call for "instruction that finds time for discussion of significant thoughtful questions" and for taking students "well beyond formal skills and critical thinking to help them through their own active learning"(pp. 41, 25).[84] The hard content for all students reform can also be seen in the national goals announced by President George Bush.[85] Of the six goals, one states: "By the year 2000, American students will leave grades four, eight, and twelve having demonstrated competency over challenging subject matter, including English, mathematics, science, history, and geography; and every school in America will ensure that all students learn to use their minds well so that they may be prepared for responsible citizenship, further learning, and productive employment in a modern economy." The hard content for all students goal is seen in the words "challenging," "learning to use their minds," and "all students."

---

[83]National Research Council, Everybody Counts (Washington, DC: National Academy Press, 1989); National Council of Teachers of Mathematics, Curriculum and Evaluation Standards for School Mathematics (Reston, VA: Author, 1989); American Association for the Advancement of Science, Science for All Americans (A Project 2061 report on Literacy Goals in Science, Mathematics, and Technology) (Washington, DC: Author, 1989).

[84]Paul Gagnon, and the Bradley Commission on History in Schools (eds.). Historical Literacy: The Case for History in American Education (New York: Macmillan, 1989); See also, National Commission on Social Studies in the Schools. Charting a Course: Social Studies for the 21st Century (Washington, DC: Author, 1989).

[85]U.S. Department of Education, National Goals for Education (Washington, DC, Author, 1990).

This is, of course, not the first curriculum reform in the history of American education, but it is certainly the most challenging. The post-Sputnik reform of the 1960s called for hard content, but only for the educationally elite. The Great Society reform of the late 1960s and 70s sought to guarantee basic skills (i.e., easy content) for all students. Thus, the recent calls for curriculum reform combine the greatest challenges of these two earlier curriculum reforms. Accomplishing the goal of hard content will not be easy; education practice will need to change substantially.

Thus far, the 1980s curriculum reform is mostly rhetoric. The needed resources and policies have not been put in place to bring about the reform.

## WHAT IS THE ROLE OF TESTING IN FUTURE REFORM?

At present, many different reforms and goals coexist while practice in schools continues largely to reflect the 1970s commitment to basic skills for all students. It is not clear what directions policy and practice will take as we approach the turn of the century. Since the role of testing is different for each reform, predicting the future of testing is equally complicated.

We are persuaded by the wisdom of the late 80s curriculum reform, which combines commitments to equity and excellence. This is in the best interest of our nation, socially and economically. We are also persuaded by much of what is called for in the restructuring reforms. Like many others, we also see a tension between the two. On the one hand, restructuring calls for local control, including local control over curriculum. On the other hand,

curriculum leadership is desperately needed. We are not optimistic that this leadership will occur spontaneously at local sites across the country. The challenge is to shift the curriculum in all states, districts, and schools in the direction of hard content for all students while at the same time empowering teachers and schools. In the next section we offer an approach to solving this puzzle; testing plays a small but important role.

## WHERE SHOULD REFORM AND TESTING GO FROM HERE?

We are convinced of the need for systemic school reform.[86] Systemic reform calls for three things. First, a systemic and rational set of policies must be put in place at all levels of the school hierarchy which support restructuring and curriculum reform. Second, current policies in education, especially at the state and district levels, must be reviewed and all policies discontinued which stand in the way of restructuring efforts, on the one hand, and the goal of hard content for all students, on the other. Third, a system of education indicators must be put in place which allows progress to be monitored over time and against the goal of hard content for all students. There are several pieces of curriculum policy that seem necessary; some involve testing.

Curriculum frameworks must be put in place which articulate the goal of

---

[86]Marshall Smith and Jennifer O'Day, "Systemic school reform." In S. Fuhrman & B. Malen (eds.), The Politics of Curriculum and Testing (Falmer Press, in press); Andrew Porter, Douglas Archbald, & Alex Tyree (op cit.); Andrew Porter, "External Standards and Good Teaching: The Pros and Cons of Telling Teachers What To Do." Educational Evaluation and Policy Analysis, vol. 27, No. 4, 1989, p. 343-356.

hard content for all students for each level of schooling and for each area of academic content. These frameworks must provide a rationale that is understandable and acceptable to educators. To give the frameworks this kind of authority, local educators will need to be heavily involved in their development. The frameworks must be specific on what areas of content are most important and that all students are to have access to worthwhile content that goes beyond basic skills. California's curriculum frameworks provide the best current example of what we believe is appropriate and necessary.[87]

To give these curriculum frameworks influence, appropriate accountability systems should be put in place. Research and experience suggest a number of characteristics for these accountability tests. They must be criterion-referenced; aligned to the curriculum frameworks, not aligned to current practice. It is unlikely we can have valid tests of hard content with only multiple choice questions. More authentic forms of assessment will be necessary.[88] They will need to be criterion-referenced to be consistent with the goal of "hard content," to operationalize the vision of the frameworks, and to assess where schools stand in delivering the desired outcomes. Norm-referencing will also be needed to compare achievement of students across

---

[87]California State Department of Education. Mathematics Framework for California Public Schools: Kindergarten Through Grade Twelve (Sacramento, CA: Author, 1985).

[88]Douglas Archbald and Fred Newmann, Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School (Reston, VA: National Association of Secondary School Principals, 1988); Harold Berlak and Fred Newmann (eds.), Toward a New Science of Educational Testing and Assessment (Buffalo, NY: State University of New York Press, 1991).

levels of income, race, and gender. This is required by the goal of "all students." These assessments will need to be constructed, administered, and judged (scored) in ways that insure that each student has an equal opportunity to demonstrate what he or she knows and can do.

The tests must be created against well specified domains of student accomplishment, so that multiple forms can be developed and a different form used at each administration. Accountability must be for achievement of the curriculum frameworks, so that education is targeted to the domains being tested and not to a specific form of the test. Multiple forms will be expensive, but absolutely essential. No single form could represent the desired outcomes of schooling. Finally, accountability testing must be done on a census basis. Each school must be held accountable for providing quality education to all of its students.

We are not certain what the consequences should be for performance. States, districts, and schools must be held accountable; students must not become the scapegoats. Still, students and parents share responsibility for learning; they must be held accountable too. Who is to be held accountable and how is a complicated problem. However accountability comes to be allocated among parents, students, schools, districts, and states, each party should be rewarded for good performance with rewards sufficient to guide curriculum and instructional practice in the direction intended by the frameworks. At the same time, not only must the accountability system <u>be</u> fair, people must believe it is fair.

Curriculum frameworks and accountability alone will not have the desired effects. More is needed than the vision of a framework and the will that accountability might produce. Know-how and belief are also needed. Many teachers lack the indepth knowledge of their subject matter that teaching hard content requires. Some teachers believe that students from low income families, minority students, or girls can't or shouldn't learn certain kinds of academic content. Today's curriculum materials are not aligned with the goal; substantially different materials will be needed. Without addressing these problems of expertise, belief, and technical support, accountability would be unfair and might make schooling worse.

The challenges to systemic reform are great. It would not be easy for educators and the public to agree upon a curriculum of hard content for all students. Schools are expected to be all things for all people. Clarity in goals also makes clear what is less important. Important outcomes will be left out, and this will make people uncomfortable. Yet effective schools require focus and systemic reform will make the focus painfully clear to all.

The challenges for professional development seem even greater. There are 2.6 million teachers in the United States, and as a group they lack the knowledge, skills, and commitment to deliver on the goal of hard content for all students. Neither preservice nor inservice teacher education, as they currently exist, is up to the challenge. The problems and probable costs are substantial. California spends approximately 1.8 percent of its education funding on staff development. This modest investment allows for a few days of

workshops each year for a fraction of the teachers.[89] No one believes this band aid approach can produce the gains in subject matter knowledge, pedagogical expertise and changes in teacher beliefs required for today's curriculum, let alone a curriculum of hard content for all students. Changes in preservice teacher education may not be as costly financially, but they will not come easily. The Holmes Group which began its reform of preservice teacher education four years ago and which is targeted on affluent and prestigious research universities, has struggled to accomplish change in practice.[90] Even the leading institutions in this group are finding resistance to change hard to overcome and their graduates all too similar to their graduates of ten years ago.

Producing useful and visionary instructional materials to support systemic reform is another challenge. Once the market for such materials is created, publishers will surely attempt to respond. But just as teachers lack the knowledge they need, the same may be true of publishers. Materials which help all students understand complicated concepts, which challenge students with novel problems and applications, are currently in short supply.

More than just textbooks are needed. Teacher-made tests must improve. The vast majority of testing is internal classroom testing, which relies

---

[89]Judith Little, et al. Staff Development in California: Public and Personal Investments, Program Patterns, and Policy Choices (San Francisco, CA: Far West Laboratory for Educational Research and Development, 1987).

[90]The Holmes Group. Tomorrow's Schools (East Lansing, MI: Author, 1990).

heavily on tests embedded in curriculum materials.[91] Teachers are not generally skilled at test construction, and curriculum embedded tests and instructional materials usually overemphasize facts and skills, since these are most easily taught and tested. High quality classroom tests aligned with state curriculum frameworks should be made available to provide teachers with both diagnostic and evaluation information. Technology could play an important role in providing instructional support to teachers and students, but this has not happened yet in any substantial way.

We believe accountability should be the last piece of systemic reform put in place. It would be unfair to hold schools and students accountable for the ambitious goal of hard content for all students without first providing the support necessary for achieving that goal. When basic skills were the goal, policy could lead with accountability, since basic skills have always been a part of the curriculum, since most teachers have knowledge and beliefs consistent with the goal, and since materials are supportive of the goal. But when the goal represents fundamental change from current practice, technical assistance must precede accountability. There is a paradox here. The primary reason for accountability has been to motivate desired practice and performance. How will shifts in practice be stimulated without leading through the heavy hand of accountability? Leadership from politicians, educators, professional societies, business and industry, and the universities is vital but may not be enough. We worry whether early commitments will be sustained

---

[91]Richard Stiggins and Nancy Bridgeford, "The Ecology of Classroom Assessment," Journal of Educational Measurement, vol. 22., No. 4, Winter, 1985, pp. 271-286.

when the full costs of the reform are realized.

There is yet a third role for testing to play in systemic reform. A system of education indicators should be created immediately to establish a baseline on student achievement and to monitor progress over time. Accountability data cannot serve this measuring purpose since they are an integral part of the reform itself. Also, accountability and other pieces of systemic reform will surely vary if not from district to district at least from state to state. Indicators at the national level will help to identify which variations are the most effective.

Most of the characteristics that we outlined above for accountability tests would also hold for education indicators. Indicators need not take a census approach, however. Still, sampling must support the kinds of comparative data necessary. The system of education indicators must provide contrasts on race, gender, and family income. Since the reform is national, so should be the indicator system. Testing need not be done at every grade in every year, but it must be done in every subject. Probably testing every other year in grades four, eight, and twelve would provide sufficient monitoring of the system. Contrasts in the performance of different types of schools would also be necessary. At least urban, rural, and suburban schools should be distinguished, since the challenges to education are substantially different across these three contexts. State contrasts should be provided, since approaches to systemic reform will vary by state.

The purpose of this indicator information is to monitor the system, not

to hold schools or students accountable. Because the stakes for specific schools or students will not be high, the need to change the tests periodically is not as great as for accountability tests. On the other hand, students must be motivated to perform well on indicator tests, so that their performance is an accurate representation of what they know and know how to do. This is a serious problem with an uncertain solution. While substantial improvements have been made to the National Assessment of Education Progress, NAEP is too limited in the breadth of its content and is not currently designed to assess the higher-order forms of knowledge reflected in the recent curriculum reform reports.

## CONCLUSION

Describing, analyzing, and predicting the roles of testing in education reform has been a challenge. Both testing and reform are complicated topics in their own right, and only partially overlapping. Establishing their boundaries and where they intersect has not always been clear-cut, and some will surely disagree with what we have decided not to include. For example, a great deal of testing happens within classrooms for instructional purposes. We have touched upon this only insofar as it relates to the need for larger systemic reforms which would require consitency between internal classroom testing and mandated testing. Others may feel we gave insufficient attention to the details of restructuring reforms, though in our analysis testing plays such a small role in these reforms that, while we agree they are important, we gave them relatively little attention. Some may wish for a more historical account, arguing that our coverage since the mid-70s is an incomplete, if not

misleading, account. We agree our account is incomplete, but we hope not misleading.

It is not just the complexities of the two topics and establishing their intersections that has made our task difficult. Research on testing and research on reform are seriously incomplete. In the case of reform, most research focuses on documenting the intentions of reform, describing federal, state, and district policy practices. More recently, research on reform has attempted to assess implementation. This work is enormously informative, but typically stops short of the classroom door. Another weakness with implementation research is that typically it starts with a policy initiative and attempts to track that initiative into practice. While such an approach is logical, it also probably overstates the quality of implementation. More research is needed which starts with practice and moves in the opposite direction, attempting to attribute that practice to policy as well as a variety of other explanations. Such an approach would describe not only implementation but also the importance of implementation relative to total variation in practice. When it comes to assessing the effects of reform, looking beyond implementation to outcomes, there is virtually no research.

Similarly, research on testing is seriously incomplete. There is a great deal of psychometric research documenting the reliability, validity, and scaling properties of various tests. This psychometric literature has made important contributions, but is not particularly helpful for our task of analyzing the role of testing in reform. More useful would be research on test use, both the amount and the type of testing done. Oft-heard claims of too

much testing and damaging effects on curriculum and instruction appear to us not to be well supported by research. At the same time, mandated testing provides less accountabilty and instructional guidance than backers of mandated testing assume. Given the small number of well-designed studies of testing and the substantial variation in kinds and purposes of tests, testing conditions, student populations, and types of potential effects worthy of concern we believe the research base remains too thin to justify strong generalizations.

The reasons for this relative lack of research on reforms and on testing is not clear. Possibly one reason is because many people think they already know the answers. Those involved in reform believe the results of the reform will be positive. Critics of testing are convinced tests force teachers into bad instructional practices, discriminate against minorities, and waste valuable instructional time. Similarly, advocates of mandated tests presume that if we just hold teachers and students accountable, performance will improve. Of course, there is enough truth to all these beliefs that they continue. Our analysis convinces us that things are more complicated and that practice might profit from better research and a more open mind.

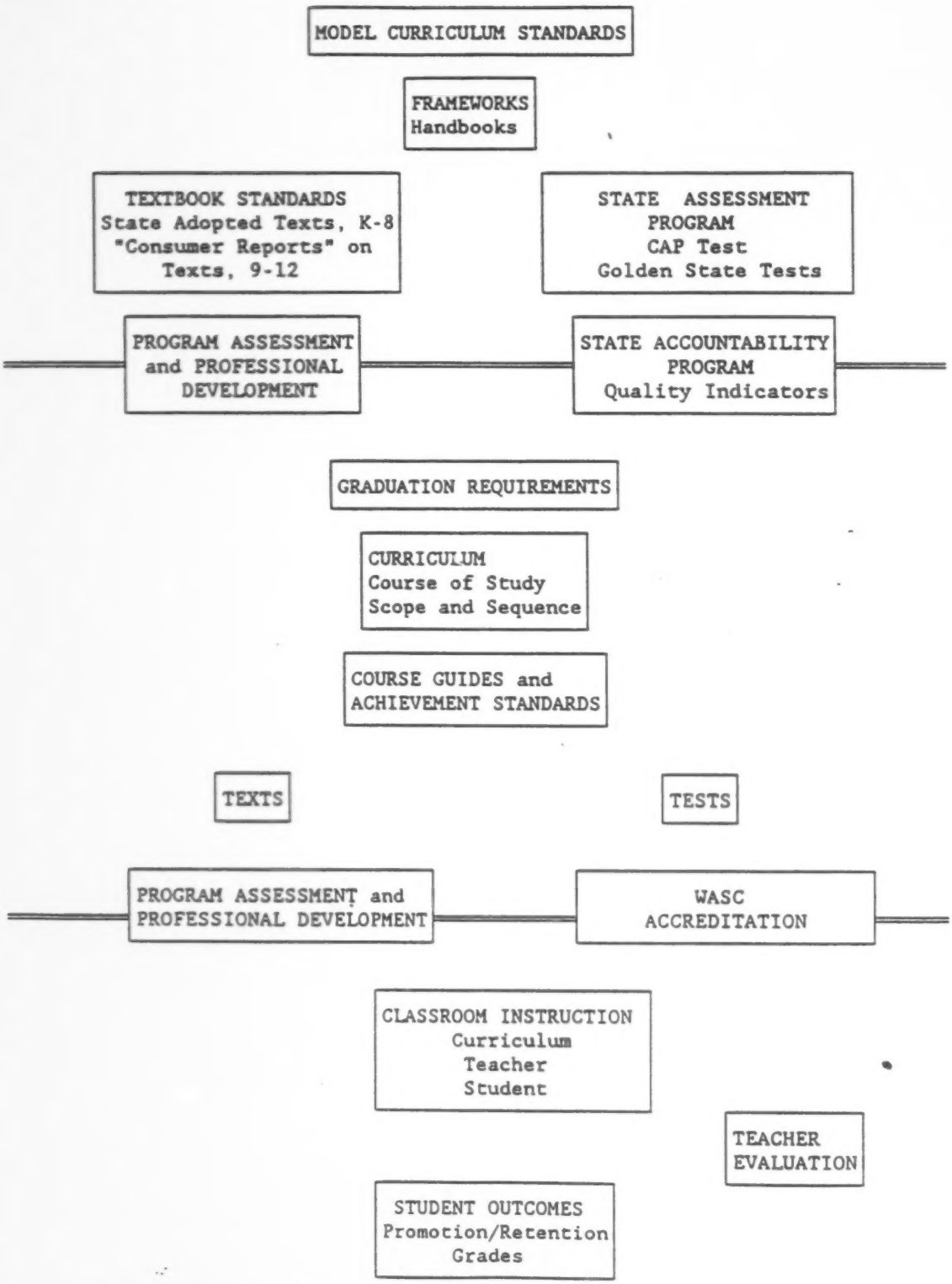
Based on the information we have been able to find, we think testing has played an important role in reform, but that it is not as potent a policy instrument, either positively or negatively, as many believe. We doubt that there would have been as many reforms as there were had there not been the worry about poor test performance. Paradoxically, at the same time as national worries about poor test performance grow, so too does the number of critics of

testing. As a nation we are clearly of two minds about tests.

It is clear American education is in serious need of reform. We need not only a substantial shift in the focus of our curriculum, but also a substantial restructuring of our schools. On the restructuring side, testing can play the role of initiator and monitor. Poor achievement helps make the case for restructuring; and monitoring student achievement over time will help determine whether restructuring has been useful. For curriculum reform, testing plays not only roles of initiator and monitor, but also "participant." Testing is needed for accountability for the goals of curriculum reform, and testing aligned with the goals of curriculum reform can help guide and support instructional practice. We are convinced, however, that curriculum reform should not begin with accountability. Rather, curriculum reform should begin with professional development and technical support in the form of better and more appropriate instructional materials.

These reforms may seem too costly. Simply putting in place an accountability system will be an attractive option, at least for politicians. Such an approach is relatively inexpensive, relatively easy to accomplish, and relatively quick. When this approach fails, as it almost surely would, schools and students will be blamed, since they are the ones held accountable. Such an approach would be unfair and almost surely have disastrous results far more costly than the comprehensive education reforms we need.

"Appendix A" California Curriculum Alignment




Source: NASSP Bulletin, Vol. 71, 1987.

AMERICAN HISTORY

7.0 The Nation at Odds, 1850-1877

- @7.1 Conflict
  - 7.1.1 Sectionalism
  - 7.1.2 Unresolved conflict
  - 7.1.3 Outcome
- @*7.2 Increasing Friction
- @7.3 Political System
  - *7.3.1 Compromise of 1850
  - *7.3.2 Appeasement
  - 7.3.3 Compromise: nature
  - 7.3.4 Unresolved issues
- @7.4 Mounting Tension
- @*7.5 Elections
- @*7.6 Sectional Differences
- @*7.7 Secession
- @7.8 Battle
  - *7.8.1 Goal
  - 7.8.2 Strategy
  - 7.8.3 Emancipation Proclamation
  - 7.8.4 Lee/Grant
- @7.9 Rebuilding
  - 7.9.1 Plans contrasted
  - 7.9.2 Amendments: purpose
  - 7.9.3 Reconstruction
  - 7.9.4 Compromise of 1877
- @7.10 Position: Blacks
- @*7.11 Cartoon
- @7.12 "Frontier"

8.0 The Gilded Age: From Agrarianism to Industrialism, 1865-1910

- @8.1 Society Transformed
  - 8.1.1 Population
  - 8.1.2 Cities
  - 8.1.3 Economy
  - 8.1.4 Frontier
  - 8.1.5 Technology
  - 8.1.6 Opportunities
  - 8.1.7 Government
  - 8.1.8 Immigration
- @*8.2 National Economy
  - 8.2.1 Inventions
  - 8.2.2 Production/ Retailing
  - 8.2.3 Big business
- @8.3 Corruption
  - 8.3.1 Reasons
  - 8.3.2 Spoils system
  - *8.3.3 Civil Service reform
- @*8.4 Labor Movement
  - 8.4.1 Big business
  - 8.4.2 Opposition
  - 8.4.3 Violence
- @8.5 Effects of Changes
  - 8.5.1 Rise of business
  - 8.5.2 Urbanization
  - 8.5.3 Careers
- @8.6 Frontier Closed
  - *8.6.1 Technology
  - *8.6.2 Opportunities
  - 8.6.3 Cities
  - 8.6.4 Violence
  - 8.6.5 Government: Indians
- @8.7 Farmers
- @8.8 Third Parties
  - *8.8.1 Reasons
  - 8.8.2 Reforms
  - 8.8.3 Decline
  - *8.8.4 Contribution
- @8.9 Laissez-Faire
  - 8.9.1 Definition
  - *8.9.2 Departure
  - *8.9.3 Interstate Commerce Act
- @*8.10 Immigration: Effects
  - *8.10.1 New wave
  - 8.10.2 Problems
  - 8.10.3 Contributions
  - 8.10.4 Cities
- @8.11 Ethnic Groups
  - 8.11.1 Native cultures
  - 8.11.2 Contemporary
  - 8.11.3 Contest
- @8.12 "Frontier"

9.0 Progressivism: A New Era

- @9.1 Reform
  - 9.1.1 Leadership
  - 9.1.2 Improvement
  - 9.1.3 Extending
  - 9.1.4 Government
  - 9.1.5 Lasting reforms

@ indicates state curriculum framework and performance standard
* indicates district minimal level skills requirements

(2) identifying the effect of rejection by American society on the Mormons.

@6.8 Students will examine the expansion of the political boundaries of the United States during this time period by:

* (1) identifying a statement that relfects the concept of "manifest destiny";

(2) comparing methods by which the United States acquired territories in the 1840's with those acquired prior to 1840;

* (3) identifying territories acquired during the 1840's from a labeled map showing the territorial growth of the United States; and

(4) identifying the impact of expansion on the issue of slavery.

@6.9 Students will constrast the "Frontier" in 1850 to that of 1820, culturally, socially, economically, and geographically.

7.0 The Nation at Odds, 1850 - 1877

@7.1 Students will examine the domestic conflict threatening American society by studying the following:

(1) the increasing sectional conflict;

(2) the failure of the political processes to resolve this conflict; and

(3) the Civil War and its results.

@*7.2 Students will identify the relationship of the tariff issue, the slavery issue, the admission of states to the Union and the Dred Scott decision to increased sectional friction.

@7.3 Students will examine the ineffectiveness of the political system in resolving sectional differences by:

* (1) identifying major provisions of the Compromise of 1850;

* (2) selecting the provisions of the Compromise of 1850 that were intended to appease the South and those that were intended to appease the North;

(3) discussing the nature of compromise as evidenced in the compromise of 1850 and the Kansas-Nebraska Act; and

(4) discussing the unresolved sectional issues despite political compromises.

@7.4 Students will identify two examples of violent outbreaks that led to mounting tension which resulted in the eventual breakdown of the political system.

@*7.5 Given two maps of the United States showing the results of the elections of 1840 and 1860, students will identify changes in the fortunes of political parties.

@*7.6 From descriptions of population size or composition, commercial activity, and economic wealth in 1860, students will select statements that characterize the North and statements that characterize the South.

@*7.7 Students will identify the relationship of the controversy over states' rights versus national interests to the secession of the Southern states.

@7.8 Students will analyze the significance of the national crisis by:

* (1) identifying Abraham Lincoln's major goal in the Civil War as preserving the nation;

(2) evaluating the differences in the overall strategy of the North and the South necessitated by the differences in their goals;

(3) identifying the importance of the Battle of Antietam in relation to foreign policy and to the issuance of the Emancipation Proclamation; and

(4) comparing the strategies of Robert E. Lee and Ulysses S. Grant.

@7.9 Students will evaluate the attempts at rebuilding the South by:

(1) contrasting the approach of the Lincoln-Johnson plan to the approach of the Radical Republican plan;

## KEY TO RECOMMENDED INSTRUCTIONAL MATERIALS

### AMERICAN HISTORY

| Performance Objective | | Instructional Implications | | Resources | |
|---|---|---|---|---|---|
| Number | Topic | Teaching Time | Tested on | County Adopted | Supplementary |
| 1.0 | The "Frontier" | 1 day | | | |
| 2.0 | Native American Culture | 4 days | | Linden, Glenn M.; Brink, Dean C.; and Huntington, Richard H. Legacy of Freedom: A History of the United States. River Forest, IL: Laidlaw Brothers., 1986. pp. 14-29. | Risjord, Norman K. and Haywoode, Terry L., People and Our Country. New York: Holt, Reinhart, and Winston, 1978. pp. 2-10. |
| 3.0 | Building a New Society | 3 days | | Linden, Brink, and Huntington, pp. 50-87. | Risjord, pp. 10-26, 29-97. |
| 4.0 | Founding a New Society | 5 days | | Linden, Brink, and Huntington, pp. 88-207. | Risjord, pp. 98-191. |
| 5.0 | Spirit of Nationalism/ Sectionalism | 4 days | | Linden, Brink, and Huntington pp. 193-258. | Risjord, pp. 192-237. |
| 6.0 | Development of an American Culture | 10 days | MLST | Linden, Brink, and Huntington pp. 197-205, 215-19, 228-305. | Risjord, pp. 238-292. |
| 7.0 | The Nation at Odds | 13 days | MLST | Linden, Brink, and Huntington pp. 219-29, 236-354. | Risjord, pp. 293-367. |

AH 28

## Correlation of State Performance Standards with District Curriculum Objectives

| State Performance Standard | District Objective Number |
|---|---|
| 7. Identify causes and effects of various changes in American historical development. | |
| The student will | |
| 7.01 explain the causes and effects of given social and intellectual movements (e.g., progressivism, abolitionist movement, and the Civil Rights movement). | 9.1, 9.2, 9.3, 9.4, 9.5, 13.1.5, 13.7 |
| 7.02 explain the causes and effects of given religious movements (e.g., Puritanism, Deism, the Great Awakening, and Transcendentalism). | 6.1.6, 6.7 |
| 7.03 explain the causes and effects of given examples of colonization and migration (e.g., immigration, manifest destiny and urbanization). | 3.2, 3.3, 6.8, 8.1.8, 8.5.2, 8.10, 8.11 |
| 7.04 explain the causes and effects of given military conflicts (e.g., Mexican War, Civil War, Spanish-American War and Korean War). | 6.8.2, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 10.1, 10.2, 13.8, 13.9 |
| 7.05 explain the causes and effects of given examples of economic systems (e.g., Plantation System, Industrial Capitalism and Utopian Socialism). | 8.2, 8.4, 8.5, 8.9 |
| 7.06 use historical reasoning to develop solutions to current human problems. | 14.12 |